TITLE: MEANS AND METHODS FOR IDENTIFYING GENES AND
PROTEINS INVOLVED IN THE PREVENTION AND/OR REPAIR
OF A REPLICATION ERROR
Inventor: Tijsterman et al.
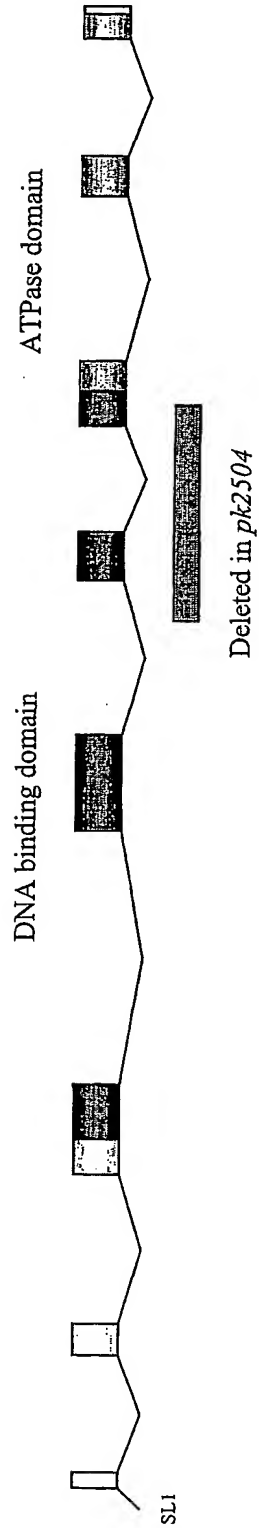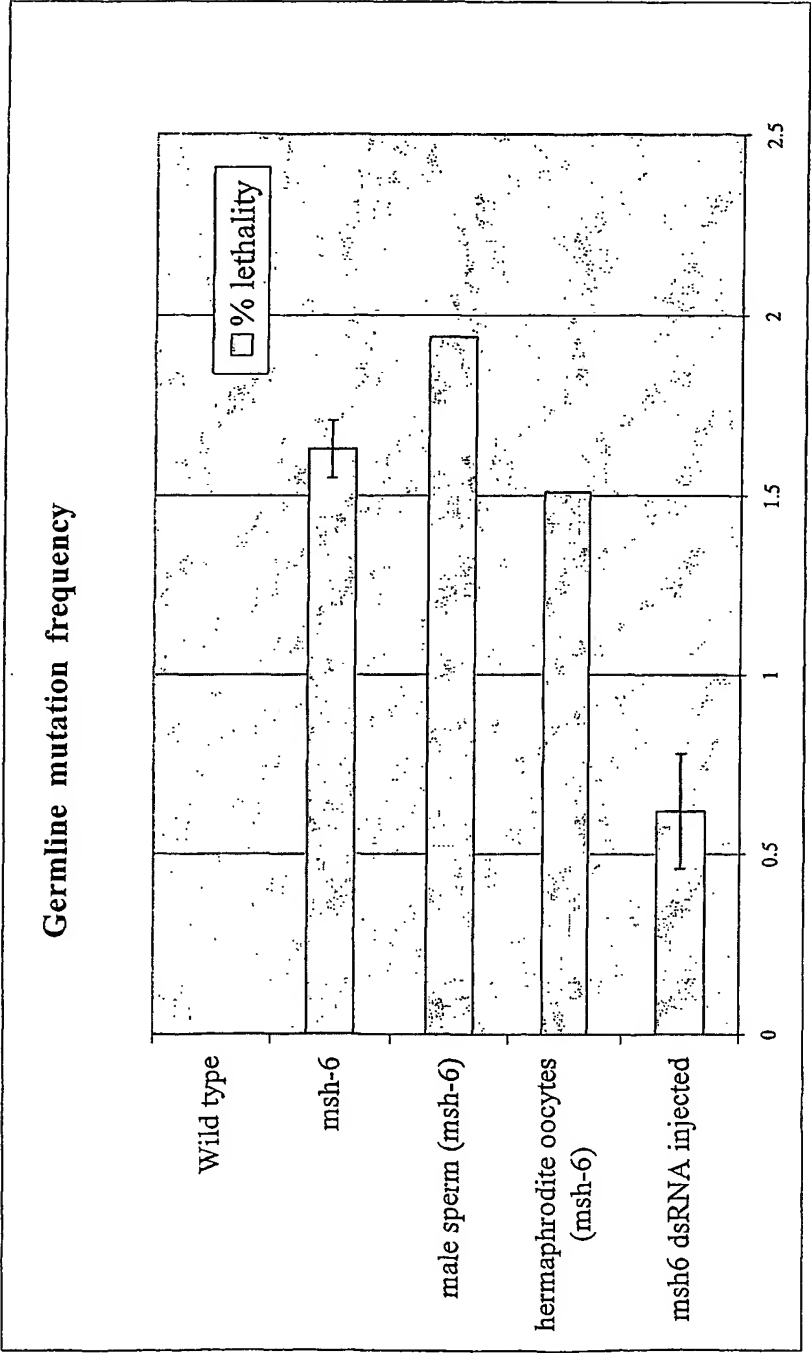Docket No.: 2183-6201US

1/7

Fig. 1A

TITLE: MEANS AND METHODS FOR IDENTIFYING GENES AND
PROTEINS INVOLVED IN THE PREVENTION AND/OR REPAIR
OF A REPLICATION ERROR
Inventor: Tijsterman et al.
Docket No.: 2183-6201US

2/7

Fig. 1B

```
C. elegans     MSKRQSSLNSFFTKTPKS EKPEEEVKEKS VE------------------------- EPKSL KNDTPKIS NDS--- 051
Human          MS RQSTLYSFF PKSP ALSDANKASARAS REGGRAAAAPEASPSPGGDAAWSEAGPGPRLARSASP PKAKNLNGGLR 077
S. cerevisiae  MAPATPKTSKTAHFENGSTSSQKKKQSSLLSFFSNQVPSG----------------- TPSK KVQKPTPATLE NTATDKI 063

C. elegans     ----------------------- ----------------------- EKKVLKRSNSKTVSSPV KTPRNASKRPKVVVCSS 081
Human          RSVAPAAPTSCDFSPGDLVWAKMEGYPWWPCLVYNHPFDGTFIREKGKSVRVHVQFFDDSPTRGWRELLKPYTGSKSKEAQKGGHFYSAKPEIL RAMQ 177
S. cerevisiae  TKNPQGG--------------------------------------- KTGKLFVDVDEDNDLTIAEETVSTVRSDIMHSQE 102

C. elegans     SEGEDDDGD----------- EDFEMKEAEHESSDESEADENASDCEVVESPESTPQSTPK RGGKKKI SKPLLAE NTP---------- -KSVKMAGKSKK 159
Human          RAEEALNKDKIKRLELAVCDEPSEPEEE BEMEVGTTYVTIKSEEDNEIESEE EVQPKTQGSRRSS RQI KKRRVIS DSESDIGGSDVEFKPDT KEEGSSDE 277
S. cerevisiae  PQSDTMLNSN------- TTEPKSTTTDEDLSSSQSRRNHKRRVNYAESDDIDSDTTFTAK RKKGKVVDSESDEDEYLPDKNDGDEDDDIADDKEDIKE 194

C. elegans     VIP ---DGEAVSMAG-- VLDKMDKIMIGEG-E RKRIVEK TTGAKNKAVELEPAER----------------------- FDHESF DFLKPDK 221
Human          I SSGVGDSESEGLNSPVK VARNRKRMVTNGSLKRKSSRKE TPSATKQATSISSETKNTLRAFSAPQNSESQAHVSGGGDDSSRPTVWYHETLEVLKEEK 377
S. cerevisiae  LAEDSCDDDDUSLAETTSKK KFSYNTSHSSSPFT RNISRDNSKKKSRPNQAPSRSYN----------- PSHSQPSATSKSSKFNKQNE ERYQVLVIE- 281

C. elegans     IRDGFKRPMSD EEYLPKTLWVEPDFHQKQTEGHRQWWTMKSQHFDTILLEKVGKEYBTYHMDAVEVVRALNIAFMRG---SYAH AGFEEHRASKIFADQIM 318
Human          RRDEHRBRPDH GDFCASTLYVFEDFLNSCTEGMRKWEQIKSQHFDTLVFMKDALIGVSELLGLVFMKG---NWAHSGFEEIAFGRYSDSLV 474
S. cerevisiae  -RDAQRRPKSDEEYDPRTLYIESSAWNKFTPFEKQYEIKSKMWDCIVFEKKGKEFBLYEKDALLANALFDLKIAGGGRANMQL AGIPEMSFEYWAAQFI 382

C. elegans     NHGYKVARIEQTPTPQML BERNQKTKTK---EK VVRREVCRVTSNGTRTYGVLDGVDLGSASSTLDPTAKH LLAIKEFHNP---ETG KS-SYGVCMIDTT 411
Human          QKGYKVARVEQTPEMMBARCRKMAHISKYDR VVRREICRIITKGTOTYSVLEGDPS-------ENYSKY LLSLKEKEED--- SSGHTRAYGVCFVDTS 564
S. cerevisiae  QMGYKVAKVDQRESMIAKQMRE-----G--S-KG IVKRELQCILTSGTLTDGDMLHSDLATFCLAIREEPGN FYNETQLDSSTIVQKL NTKIEGAAFIDTA 475

C. elegans     TAHIRIGQEEDDYRSQLRTLLANVIVVQAIVERGSISSTTKSIING-ILFSVPV EHLLPKKQEMTAEDVVRI VSNEDYYGSDAS------EW EEVLKGML 505
Human          LGK FFIGQSDDRHCSRERTTVAHYPPVQVLFEKGNLSKETKTILKS-SLSCSLQEGLIPGSQEWDASKTLRT LLEEEYFREKLSDGIGVML EQVLKGMT 663
S. cerevisiae  TGELQMLEEEDDSECTKLDTLMSQVRPMEVVMERNNLSTLANKIVKFNSAPNAIF NEVKAGEFYDCDKTYAE IISSEYFSTEED------W EEVLKSYY 569

C. elegans     EDS-SILPKPSTDW QLALSAFGAIFWYLRDSLIDVIMLSMRNVTIKN---SNSMENDQKKEK IDWNGKNLILDGTAJENLNIIVPNGRD-SHLTSLYYVIN 600
Human          SESDSIGLTPGEKS ELALSALCGCVFYLKKCLTDQELLSMANFEEYIPLDSDTVSTTRSGAI FTKAYQRMVLDAVTLNNLEIFLNGTNGSTEGTLLERVD 763
S. cerevisiae  DTG---------K KVGESAFGGLLYYLKWLKLDKNLISYKNIKEYD--------------- FVKSQHSMVLDGITLONFEIFSNSFDGSDKGTLFKLFN 644

C. elegans     KCSTPFFGPRLLRSFILQETCDPKK LEQRQKAIKWLVSPDASSEMTTATAT LKRIPDLDRLLQKIHTIGLKYRSEKHPDSRAIFFDTIKTNQ KNIAELLAA 699
Human          TCHTPFCGKRLLKQFLCAPLCNHYA INDRLDAIEDLMV-- VPDKISEVVEL LKRLPDLERILSKIHNVGSPLKSQNHPDSRAIMYEETTYSKK KIIDFLSA 860
S. cerevisiae  RAI TPMCKRMMK KIMHPLLRKND IESRLDSVDSLLQD--ITLREQLEIT FSKLPDLERMLARIHSR-------------------TI KVKDFEKV 731

C. elegans     IDGEKLCN KLRKEYI KVQKEGEGCELLDELLGNEQE------MEE VDENIYYFEKMGDRSTAMKDGK IVENAGCEEYDEALNRVKBALNEDNDYKDSVA K 795
Human          LEGGKVMCKLIGIME EVA-DGFKSKILKQVISLQTKNPEGRFPD LTVELNRWDTAGDHEKARKTGLITEKAGFDSDYDQALADIRENEQSULEYLEKQR N 960
S. cerevisiae  ITAETII ELQDSLK NNDLKGDVSKYISSFP----------EG LVEAVKSWTNAGERQKAINENIIVEQRGFDIEFEKSMDRIQELEDEDMBILMTYR K 808

C. elegans     KYSES-IIKEVDSGKVKYLLEMFENTKVS-------SS FELKSRRKGFIRYSTPDSEQLVAA LDAVEKEKSKLGDDATRRRVFEQEGHKN-PIWLETVKLVSS 887
Human          RIGGCRTIVYWGI GRNRMQLEIBENFTTR---NLPEE YELKSTKKGCKRYWTKTIEKKLAN LINAEERRDVSLKDCMRRLFYNDKNY-KDWQSAVECIAV 1056
S. cerevisiae  QFKGSNIQYKDSGKEITIEIBISATKN----VPSN WVQMAANKTYKRYYSDEVRAIARS MAEAKEIHKTLEE DLKNRLCQKEDAHYNTIAMPTIQAISN 904

C. elegans     FDVLTSLALFAKSSPFEM CMEEDFNATDP-------- YLIVDKGV HPCLALQSRNE--- VTQTTS FIANSTTMGASEAAVM ELTGPNMGGKSTIMRQTAV 977
Human          LDVLLCLANYSRGGDGPM CREVILLPEDTPP------- FLELKGSR HPCITKTFFGDD-F IPNDILIGCEEEEQENGKAYCV LVTGPNMGGKSTLMRQAGL 1149
S. cerevisiae  IDCLLAITRTSEYLGAPS CRETIVDEVDSKTNTQLNG FLKFKSLR HPCFNLGATTAKDF IPNDIELGKEQPRLG-------- ILIGANAAGKSTILRMACI 997

C. elegans     LAILDHIGSMVPAFSMR LTPIDRIFTRIGANGRIMCGESTFFIELKETDIMLKNATKHSLLLVDELGRG TSTFDGTAIASAVLQKISDDLACRTFSSTHY 1077
Human          LAVMAQMGCY VPAEVCR LTPIDRVFTPLCASDRIMSGESTFFVELSETASIDMHATAHSLVLVDELGRG TATFDGTAIANAWVKELAETIKCRTLSSPHY 1249
S. cerevisiae  AVIMAQMGCY VECESAV LTPIDR IMTRLGANFINIMQEKSTFFVELAETKKIDMATNRSLLVVDELGRG GSSSDCFAIAESVLHHVATHIQSLGFGATHY 1097

C. elegans     HSICDS FTMHPNVRLAHMKCVVDKENNEDP TMEDVTFLYELESGICPK SYEFYAAKLAG IDHQVVRNAYLE SNKFASNLI IDPKIRHLV ECARDDNFDVG 1177
Human          HSLVEDYSQNVAVRLGHMACMVEN-ECEDP SQETITFLYKFIKGACPK SYGENARLANLPEEVIQKGHRKAREFEK--- MNQSLRLFR EVCLAS----- 1340
S. cerevisiae  GTLASS FKHHPQVRPLMMSILVDE------ATRNVTFLYKMLEGQSEGSFGMHVASMCGISKEIIDNAQIAADNLEHTSR LVKERDLAA NNLNGEVVSVP 1191

C. elegans     ELKRMIEAI 1186
Human          ---------
S. cerevisiae  GGLQSDFVRIAYGDGLKNTKLGSGEGVLNYDWNIKRNVLKSLFSIIDDLQS 1242
```

TITLE: MEANS AND METHODS FOR IDENTIFYING GENES AND
PROTEINS INVOLVED IN THE PREVENTION AND/OR REPAIR
OF A REPLICATION ERROR
Inventor: Tijsterman et al.
Docket No.: 2183-6201US

3/7

Spontaneous germline mutation frequency in wildtype *C. elegans*, *msh-6* genetic mutants and wildtype *C. elegans* exposed to *msh-6* dsRNA.
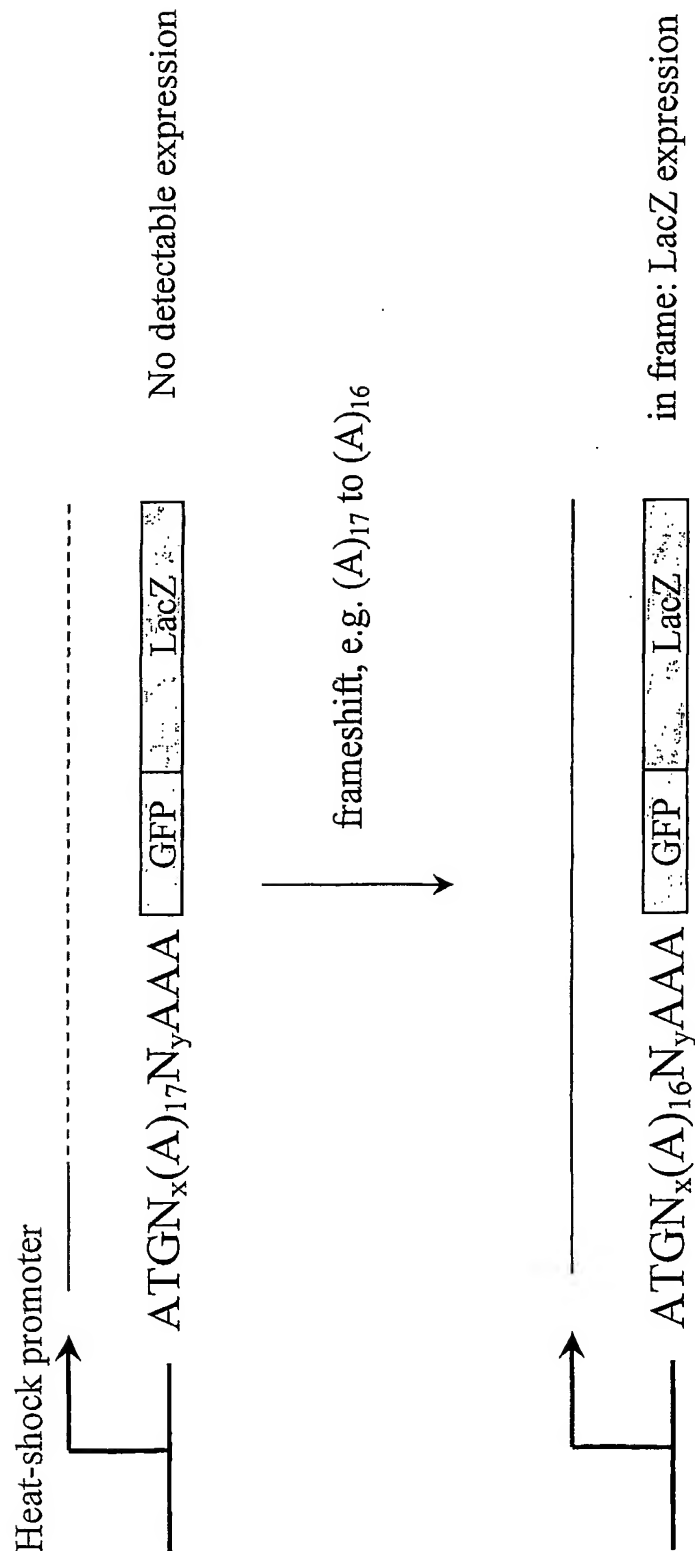
Fig. 2

BEST AVAILABLE COPY

TITLE: MEANS AND METHODS FOR IDENTIFYING GENES AND
PROTEINS INVOLVED IN THE PREVENTION AND/OR REPAIR
OF A REPLICATION ERROR
Inventor: Tijsterman et al.
Docket No.: 2183-6201US

4/7

Heat-shock promoter

$ATGN_x(A)_{17}N_yAAA$ | GFP | LacZ

No detectable expression

frameshift, e.g. $(A)_{17}$ to $(A)_{16}$

$ATGN_x(A)_{16}N_yAAA$ | GFP | LacZ

in frame: LacZ expression

Fig. 3: Outline of the principle to detect somatic repeat instability

TITLE: MEANS AND METHODS FOR IDENTIFYING GENES AND
PROTEINS INVOLVED IN THE PREVENTION AND/OR REPAIR
OF A REPLICATION ERROR
Inventor: Tijsterman et al.
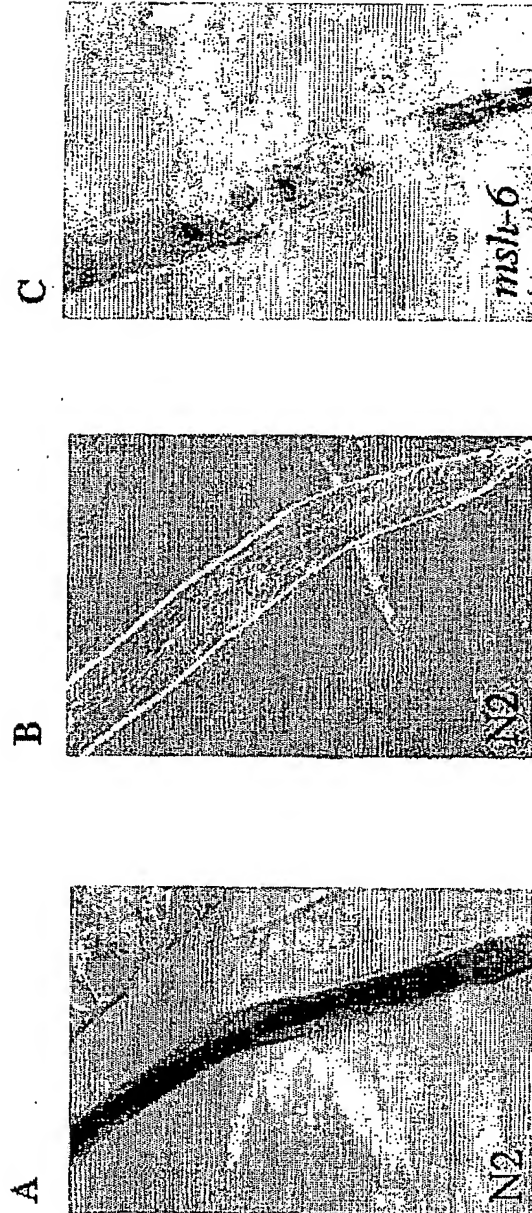Docket No.: 2183-6201US

5/7

Figure 4: A) Wildtype *C.elegans* containing the in-frame construct, B) the +1 out of frame construct. C) Genetic *msh-6* mutants that contain the +1 out of frame construct display LacZ expression.
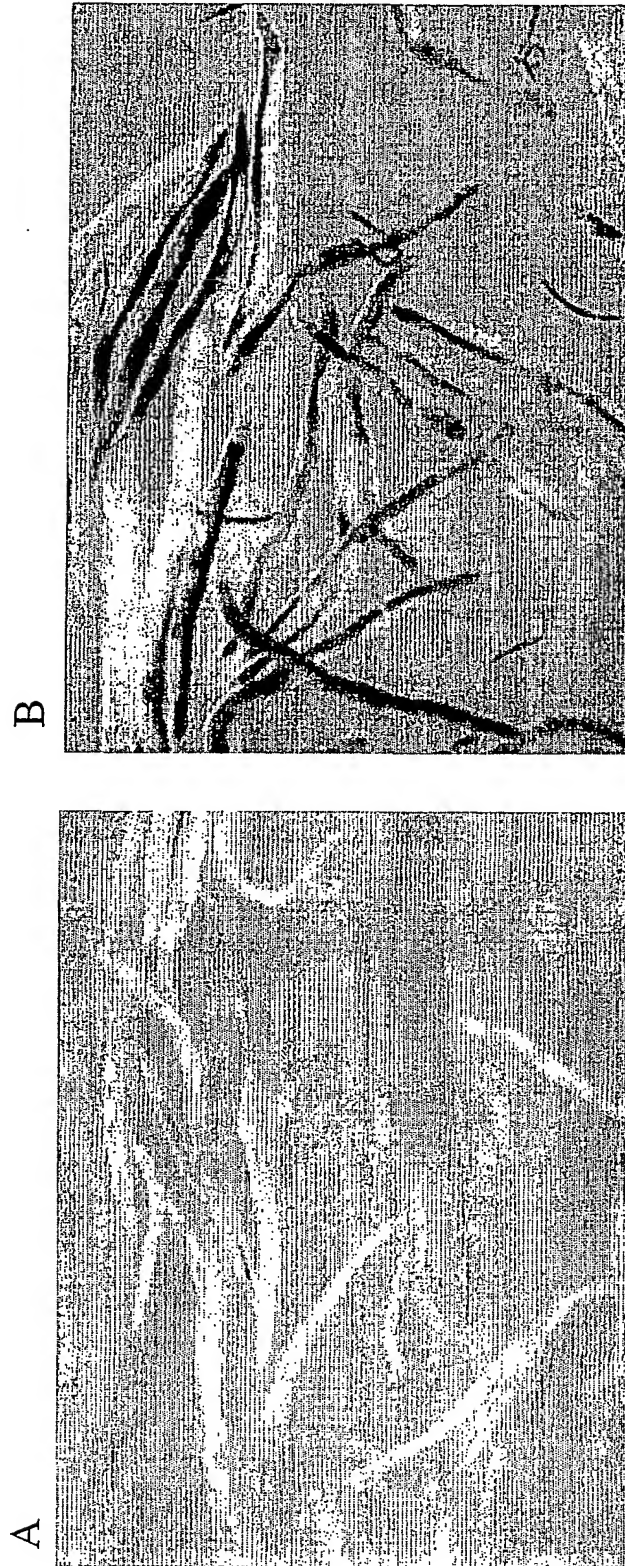
TITLE: MEANS AND METHODS FOR IDENTIFYING GENES AND
PROTEINS INVOLVED IN THE PREVENTION AND/OR REPAIR
OF A REPLICATION ERROR
Inventor: Tijsterman et al.
Docket No.: 2183-6201US

6/7

Fig 5: *C. elegans* populations fed on *E.coli* that produce dsRNA homologues to the *C. elegans* genes *unc-22* (A) and *msh-6* (B)

TITLE: MEANS AND METHODS FOR IDENTIFYING GENES AND
PROTEINS INVOLVED IN THE PREVENTION AND/OR REPAIR
OF A REPLICATION ERROR
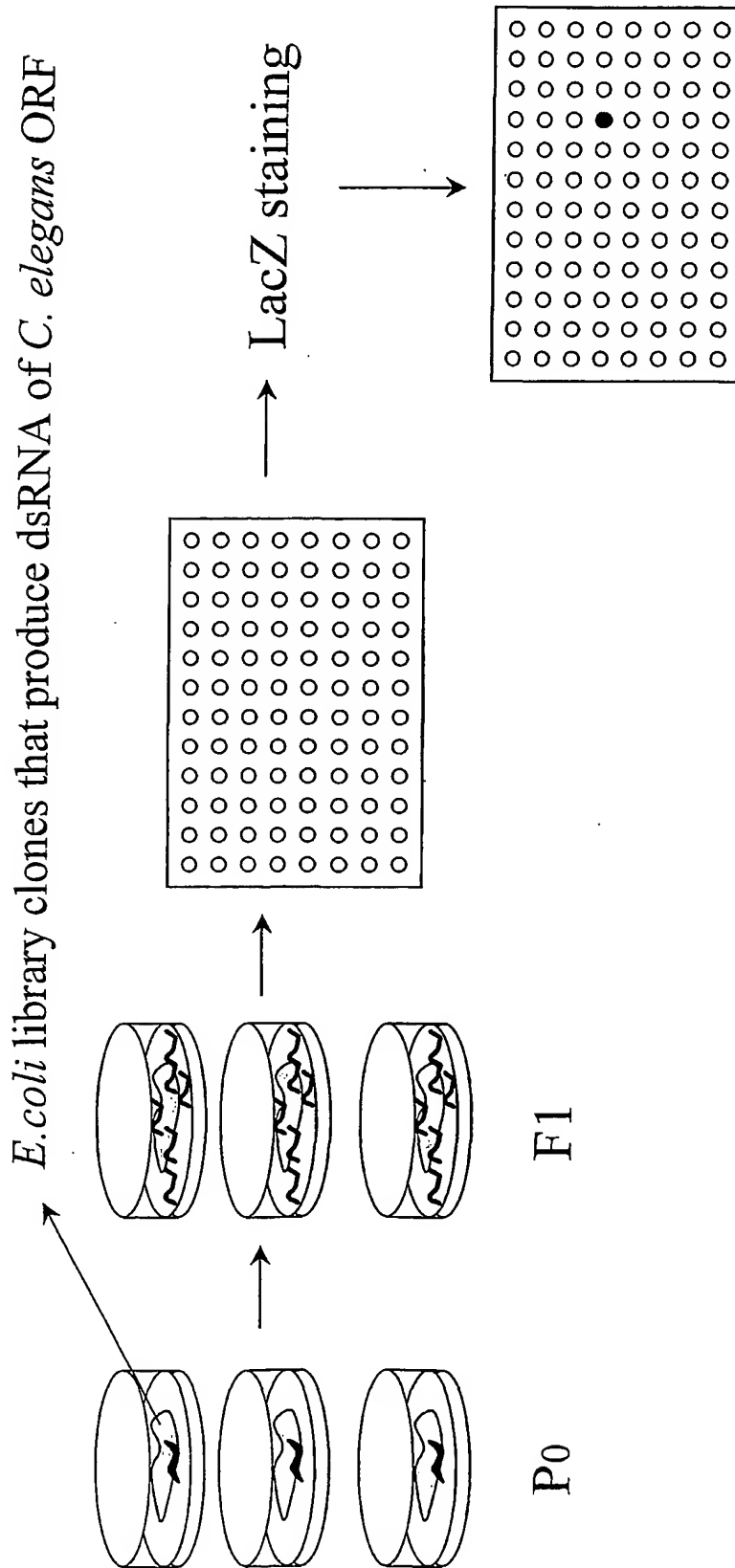Inventor: Tijsterman et al.
Docket No.: 2183-6201US

7/7

Fig. 6: Schematic representation of the high throughput RNAi based screens to identify novel mutator loci: Individual animals are fed on dsRNA producing bacteria, the progeny is collected and assayed for beta-galactosidase activity.